5. QUALITY ASSURANCE & QUALITY CONTROL

Although often used interchangeably, "Quality Control (QC)" and "Quality Assurance (QA)" are two different, but related, ideas. The two ideas were defined very well by Donald E. King in Credibility: The Consequence of Quality Assurance, from the May 1982 Data Quality Report Series, Ontario (Canada) Ministry of the Environment.

> **Quality assurance** is described as a management function which rests on the documentation and establishment of quality control protocols, and on the evaluation and summarization of their outcomes.
>
> **Quality control** is a technical, operational function which investigates and confirms the proper conduct of all those procedural components necessary to a successful conclusion.

In the context of laboratory certification, quality control is a variety of techniques that the sampler and analyst perform to verify that the sampling and analytical protocols meet the desired goals for data quality. Quality control functions help to ensure data validity and traceability. Quality assurance is the system for checking and ensuring that quality control criteria are met and what actions are taken in the event they are exceeded. The Laboratory Certification and Registration Program requires each laboratory to establish a quality assurance manual which sets up and governs the necessary quality control samples and protocols.

Using a specific example, a quality control process is the requirement to establish quality control limits for the evaluation of matrix spikes and replicates. Let's assume Acme Laboratory has established quality control limits for phosphorus matrix spikes, and those limits are -21% to 300% recovery. We all agree that it is possible to statistically generate limits such as these. And statistics is a mathematical science. And science is not wrong. Therefore, Acme Laboratory has certainly fulfilled its quality control requirement. Right?

Anything wrong with this picture? Most operators would agree that -21 to 300% is not a very good set of quality control limits. Quality Assurance would agree. The role of Quality Assurance in the laboratory is to ensure that the quality control program will guide the laboratory towards generating data which meets its goals for accuracy and precision. In the case of Acme Laboratory, **this lab could actually forget to spike the samples, obtain a recovery of 0%, and pass muster**! The role of the Quality Assurance program then, is to continually review all aspects of the quality control program and make adjustments— or initiate corrective action— as needed to achieve the laboratory's data quality goals. An appropriate series of corrective actions, in this scenario, is to:

- review quality control limit calculations for obvious errors,
- test the data for outliers, exclude any that are identified as such, and re-calculate control limits,
- review matrix spike preparation procedures to determine if any errors were made,
- contact any area laboratory which performs this analysis—or the State Laboratory of Hygiene—and use their current quality control limits on an interim basis.

To better illustrate the distinction between QC and QA, consider the following examples:

Laboratory Activity	QC Analysis	QA Analysis
The analyst checks the balance, before using it, with a 100 mg Class "S" weight. He records a weight of 83 mg in the logbook.	☑ This is certainly a QC function which must be done.	☒ The QA failure, in this situation, is not recognizing that a difference of 17 mg (17%) from expected is unacceptable. The balance calibration should be considered suspect.
Before doing BOD analysis, the analyst records the pressure from the laboratory's barometer. It's a nice sunny June day, and the barometer reads 28.4 inches (of mercury).	☑ Award points for the fact that the lab is monitoring barometric pressure...as required by method.	☒ A working QA system would have identified that a normal barometric pressure under a summer high pressure system would measure about 29.9 to 30.3 inches.
The laboratory has not ever analyzed matrix spikes for either ammonia or phosphorus.	☒ Matrix spikes are a required QC program element for these analyses.	☒ It follows that no QC means there can be no QA either.
The laboratory has established phosphorus matrix spike control limits of -275% to +3047%	☑ Give the lab credit for establishing QC limits, **but**....	☒ As the saying goes, these are limits you could "drive a Mac truck through". The QA program is responsible for ensuring that control limits used will ensure data quality. Here, you could forget to spike the sample and still pass QC!
The laboratory keeps a logbook documenting oven temperatures for TSS. Records indicate consistent temperatures of about 180 ± 1°C.	👍 👎 This one's a toss-up. Documentation is a critical part of any QC program, however...	☒ Wrong test! 180 °C is great for TDS, but for TSS, it must be 103-105°C. A QA review of the SOP might have caught this.

QUALITY CONTROL MEASURES

Lab Facility Cleanliness
The laboratory is kept in a clean and orderly condition at all times. The room temperature is maintained as constant as possible. Care is taken to maintain air quality.

Personnel Training
Initial training for new staff is a priority. Regular continuing training is also provided to ensure competence and maintenance of analytical skills. Operator certification in the laboratory subclass is required of all lab personnel working at this plant. In addition, the principal lab analyst/technician has attended the vocational classes associated with wastewater laboratory analysis. The lab certification code (NR 149) does not require operator certification in the lab subclass. The operator certification code (NR 114) requires the operator in charge to hold a valid certificate for the appropriate subclasses and grade.

Equipment Maintenance
A file is kept for each piece of equipment in the lab. Each file contains the owner's manual, a preventative maintenance schedule, and records of all maintenance and repairs performed including the exact nature of the problem, the date of the repair, what was done, who did it, and the cost. To determine if an instrument malfunction affected analysis results, the dates of breakdown and subsequent repair are considered particularly important. The analytical balance is serviced annually. The DO and ammonia probe membranes are replaced every two to four weeks or more frequently if readings become erratic.

Analytical Reagents
Only analytical grade reagents are used. Labels on all chemical reagents are marked with the date received and opened. All reagents and standards are labeled with the date they were prepared. Chemicals are stored out of direct sunlight and refrigerated if necessary to prevent deterioration. Care is exercised to prevent cross-contamination of all reagents. Solutions which have been opened for over a year are discarded and ordered fresh. Shelf life dates are closely monitored. For reagents mixed in the lab, shelf life recommendations provided within the analytical method are followed. Care is taken to store all chemicals in a safe manner. Flammable solvents are kept in a fire proof cabinet. Strong acids and bases are stored in separate cabinets.

Reagent Water Quality
Reagent grade water is produced in the lab using a Barnstead Model A1015 tin lined still. Water used for ammonia measurements is also passed through a mixed-bed ion exchange column (Barnstead Bantam Deionizer). Only freshly prepared reagent water is used for ammonia testing to prevent the water from picking up ammonia from the air. Dilution water for BOD analyses is stored in glass carboys stoppered with clean cotton plugs. Reagent water for tests other than ammonia and BOD is stored in tightly stoppered glass carboys.

Labware Cleaning
After each use, glassware is washed with detergent, rinsed with tap water, rinsed with distilled water, allowed to dry, and stored in a cabinet. The appropriate glassware cleaning procedures depend on the analysis to be performed. Glassware for Phosphorus tests is washed with a non-phosphate detergent, acid-washed after each use, and segregated from other lab glassware. BOD bottles are always stored dry. The siphon tube used for BOD analyses is cleaned monthly with a bleach solution (25 ml Chlorox / L water) and rinsed thoroughly. All reagent water carboys are cleaned monthly with dilute hydrochloric acid.

Instrument Calibration
The pH meter, DO meter, and ammonia selective electrode are calibrated each day they are used. If these instruments are used over the course of a day, calibration checks are repeated every 2 hours. The temperature of the BOD incubator and sample storage refrigerator are measured on thermometers with their bulbs immersed in water. Each day the incubator or storage refrigerator are used, their temperatures are recorded on log sheets taped to the equipment. The temperatures of the solids drying oven, muffle furnace, and fecal coliform incubation bath are recorded on the bench sheets when they are used for analyses. If temperatures are outside of the required range, the thermostats are adjusted and the adjustments are noted on the log sheets.

Thermometers used in the lab to measure the temperature of the BOD and fecal coliform incubators are calibrated annually against a thermometer traceable to an NIST (National Institute of Standards and Technology formerly National Bureau of Standards, NBS) certified thermometer (borrowed from a neighboring wastewater treatment lab). Both thermometers are tagged with their correction factors, and the appropriate factor is applied when documenting any temperatures. If the mercury column in the thermometer splits, it will not provide an accurate measurement, and must no longer be used.

Thermometers used to record the temperature of influent and effluent 24-hour composite samplers are factory certified traceable to NIST and are sealed in a clear glass bottle filled with ethylene glycol. Each thermometer has a unique serial number and a certificate of NIST traceability.

The analytical balance is zeroed daily and the calibration checked at least monthly with two class - S weights, one in the milligram and one in the gram range. The weights used and their values are noted in a log book by the balance table. If the balance calibration is off by more than 0.5 mg, the balance requires service by a manufacturer's representative.

Quality Control Analyses

Routine analysis of blanks, replicates, standards and spikes are performed according to the frequency shown in Table 4. Blanks are subjected to the entire analytical procedure including the addition of preservatives. Results of blank analyses are treated in the manner specified by the method. Data from results of replicate and spike analyses are treated in the manner specified in the following section. Results of the known standard (glucose/glutamic acid) for BOD must be 198 ± 30.5 mg/L. Records of all of quality control analyses are kept on daily bench sheets and in a separate quality control log book.

TABLE 4 - QUALITY CONTROL SAMPLE FREQUENCY

TEST	CALIBRATION	KNOWN STANDARDS	BLANKS	REPLICATES	SPIKES
BOD	Calibrate DO meter on each analysis day.	1 GGA std weekly	Each analysis day	1 per 20 samples	NA
TSS	❑ Annually: Balance checked by service representative ❑ Monthly: Check calibration with (2) S class weights (1 g, 10 mg)	NA	NA	1 per 20 samples	NA
Ammonia-Nitrogen	Each analysis day. Alternatively, establish a calibration curve and verify its validity each day with a known standard.	❑ Daily if full calibration curve not analyzed. ❑ after 20 samples in a day	Each analysis day	1 per 20 samples	1 per 20 samples
Total Phosphorus	Each analysis day Alternatively, establish a calibration curve and verify its validity each day with a known standard.	❑ Daily if full calibration curve not analyzed. ❑ after 20 samples in a day	Each analysis day	1 per 20 samples	1 per 20 samples
pH	Each analysis day	NA	NA	1 per 20 samples	NA
Chlorine Residual	Initially, and at time of reagent change	1 per 10 analysis days	Each analysis day	1 per 20 samples	NA
Fecal Coliforms	NA	NA	NA	1 per 20 samples	NA

Blind Standards

A blind standard is a sample obtained from an external source, whose validated concentration is known to the facility supervisor, but is unknown to the lab analyst/technician. These types of samples are important indicators of quality feedback on performance is provided. This offers the lab technician an opportunity to correct any analytical problems in a timely manner. Blind standards are analyzed every four months as a check on analytical performance. Appropriate measures are taken to investigate problems when the result of a blind standard analysis is inconsistent with past results.

The Tree City facility obtains blind standards from the State Laboratory of Hygiene or Analytical Products Group (APG).

Unlike reference samples, whose results are not received for some time after the testing is performed, these samples can be used at any time there is concern about the control of a specific analysis. Having immediate access to the validated concentrations allows the analyst/technician to take immediate action to identify and correct the problem.

Known Standards
These standards are prepared by or acquired by the lab, with a validated concentration of the contained analyze. They are used to verify the accuracy of the system.

Reference Samples
A reference sample is a standard, obtained from approved external sources, whose concentration is unknown to the laboratory. For many tests at least one set of reference samples must be analyzed every year to renew the lab's registration. At least one set of acceptable results must be obtained. Follow-up reference samples are required to be analyzed if the provider acceptance limits are exceeded.

The State Lab of Hygiene ships reference samples several times a year. For each study, the laboratory will have to analyze and report results by a deadline. For the State Lab of Hygiene reference samples, the laboratory will receive a final report with the true values and acceptable ranges 30 days after the results deadline. If all the results are acceptable, the laboratory does not need to do anything else since results are electronically loaded into the lab certification computer system. If one or more failures occur, then the lab will automatically receive the next round of reference samples from the State Lab of Hygiene.

The Tree City facility obtains reference samples from the State Laboratory of Hygiene.

CALIBRATION

Though much attention is given to the generation and evaluation of known standards, reference samples, blind standards, and matrix spikes in ensuring analytical accuracy, perhaps the most overlooked aspect is calibration. Analysts are provided with a wide array of options, including instrument pre-programmed calibrations, hand-drawn calibration curves, the use of sophisticated scientific calculators, and even computer software. Unfortunately, the advantages and disadvantages of these various techniques are rarely discussed.

The calibration process entails establishing a relationship between instrument response and concentration of analyte. The following "rules" should be applied to this process to ensure the generation of accurate data.

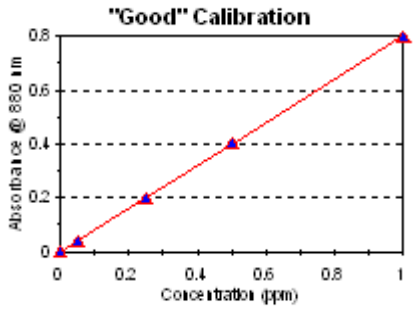
- **Use an appropriate number of standards.**
  Calibrations must be constructed using at least 3 standards and a blank.

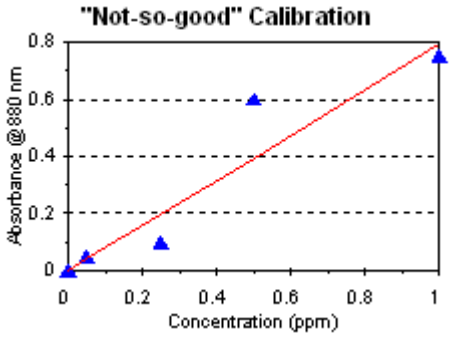
- **Keep it simple: go linear.**
  The majority of the instruments used in the environmental laboratory relate response to concentration linearly, or transform the primary signal to produce a linear output. Although deviations from linearity are encountered in the analytical range, they are more common at the extremes, where detector saturation or insensitivity are the culprits. The simplest means to generate a calibration curve is linear regression; the most defensible way to do this is to establish concentration as the **x**-axis and response to the **y**-axis.

- **Know when to include a zero.**
  Unless you are using a calibration algorithm that fits points exactly, you will have to decide what to do about blank signals. A good rule of thumb to follow is: if you can adjust your instrument to read zero in the presence of a blank, then include a zero point in your calibration curve. Including a zero is generally appropriate for colorimetric procedures which use a spectrophotometer to measure response.

- **Do not <u>force</u> curves through zero.**
  Doing this manipulates the data mathematically to obtain a y-intercept equal to zero. As a result, we lose valuable information about the lower limits of the analytical signal, and a good estimator of the limit of detection on the calibration day. Even when it is justified to include a zero point in the calibration curve, forcing the intercept to read zero is not. That being said, it must be noted that the Standard Methods procedures for phosphorus do instruct the user to, "*Plot absorbance vs. phosphate concentration to give a straight line passing through the origin.*" Consequently, while scientifically this practice should be discouraged, it remains an allowable option for phosphorus calibrations.

- **Evaluate the accuracy of your calibration**
  Consider the following calibration graphs. Most people would agree that the graph on the right is not a





  very good calibration. Unfortunately, little advice beyond, essentially, "plot concentration vs. response" is offered by most analytical methods. Consequently, these types of curves are occasionally encountered. Take note that a line "of best fit" drawn through a set of calibration points is just that. Without a means of evaluating the resultant line, data quality can suffer.

  One measure of a particular curve's validity is the correlation coefficient. A correlation coefficient of at least 0.995 generally indicates acceptable characterization of the curve. If this degree of correlation is not obtained, the reason for the lack of linearity should be investigated, any necessary corrective action taken, and a new calibration curve must be constructed. Many inexpensive scientific calculators provide the correlation coefficient of a set of data with a single keystroke.

  Another way of evaluating a calibration curve is to utilize the calibration equation (e.g., slope and intercept) to convert the response obtained for the calibration standards into concentration. This "predicted" concentration should agree reasonably well with the known, or "true" concentration of the standards. Opinions differ as to what constitutes "reasonable" agreement, however, if the predicted concentration is generally within 5-10% of the "true" concentration, the calibration is acceptable.

- **Define your calibration range properly.**
  You would not use a telescope to examine a cell, or a microscope to observe lunar craters. Similarly, the calibration range should be appropriate for the samples being analyzed (i.e. don't calibrate from 1 to 5 mg/L when all the samples fall between 0.05 and 0.5 mg/L). For low level work, it is best to choose points above, but near detection limits. For high level work, defining the upper limit of the calibration range is more important. The most accurate results are obtained when the signals of unknowns (i.e., samples) are close to those found in the knowns used to establish the calibration curve.

  A good calibration curve is like a well-maintained highway: it has legible signs and evenly spaced markers. All curves should be accompanied by the equations or coefficients that define them and should be generated using, as much as possible, evenly distributed points, the more (at least three), the better. Always use the <u>full</u> calibration curve to quantitate samples. **Never** use a daily calibration verification standard to quantitate samples.

Pre-programmed Calibrations
A number of commercially available spectrophotometers offer “pre-programmed” calibration curves for many of the routine wastewater tests, including chlorine residual, and phosphorus. The use of pre-programmed calibrations is unacceptable. The laboratory must generate its own standard curve. A manufacturer's claim that their method is approved or acceptable does not mean that the approval extends to pre-programmed calibrations. When the EPA extends “approval” to one of these manufacturers that their particular technique is “equivalent” to a referenced EPA method, the approval is granted on the basis of no significant difference in the stoichiometry or chemistry of the procedure.

Hand-drawn Calibration Curves
Many laboratories use a calibration curve constructed manually by plotting the concentration of phosphorus on the x-axis and absorbance on the y-axis. A straight line which best fits the data points is then drawn, and sample concentrations are determined using the "best fit" line to convert absorbance into concentration.

The laboratory certification and registration program discourages this practice because there is significant variability in both how the scale of the graph is constructed, and how any individual draws the "best fit" line through the calibration data points. This degree of variability makes it difficult to trace your results as they appear on the discharge monitoring report (DMR) back to the raw data. The best fit line becomes the one YOU drew, and not necessarily the one that best represents the relationship between concentration and instrument response. When a certification officer comes to your laboratory, one part of the audit process will be to verify that the absorbance for a particular sample indeed relates to the concentration reported on the DMR. In addition to the potential for the certification officer to read a different concentration from the curve, the auditor may feel that the line has not been drawn accurately (i.e., it is not the "best fit" line). Traceability of results is a critical requirement of laboratory record-keeping practices, and is described in section NR 149.06, Wis. Adm. Code.

Scientific Calculators
Sources of variability can be eliminated if a standard procedure is used to generate a calibration function. One of the most widely recognized means of achieving this is the use of a linear regression. It is likely that you will have already heard your certification officer discussing linear regressions. Linear regression equations can be generated with an inexpensive scientific calculator and most spreadsheet programs. Linear regression is a statistical procedure that will produce consistent equations for a "best-fit" line, eliminating questions or bias in the validity of a hand-drawn line. The DNR is working to provide laboratories with more assistance in this area. Look forward to more information about this in future issues of *LabNotes*.

Computer Software
Finally, a wide variety of PC-based spreadsheet programs are available.

DETERMINING THE LIMIT OF DETECTION (LOD)

A critical aspect of any analysis is the point at which a sample result can be distinguished from a blank. This concentration has historically been termed the method detection limit (MDL), but is now known as the Limit of Detection (LOD). As analytical results approach the LOD, precision is generally reduced, so it is important to understand what the LOD is when measuring, reporting, or evaluating the results for any sample.

Section NR 149.11 (5) requires laboratories to determine the LOD "*for each analyte reported by a laboratory in accordance with a method specified by the Department*". Although the conventional method for determining LODs is based on the analysis of spiked samples prepared using laboratory reagent water, the Department may also require that the limit of detection be determined for a specific matrix. For wastewater laboratories that are only analyzing their own effluent, it is a good practice, although not a requirement, to determine their LOD using their effluent as a matrix. The most common, and simplest, procedure for determining the LOD can be summarized as:

- Determine a spike concentration, which approximates the LOD
- Prepare 7 or 8 replicate of reagent water spiked at an appropriate level
- Analyze the replicate spikes
- Calculate the LOD
- Perform the "5-point" check of the LOD
- Repeat the LOD determination as appropriate

Choosing an Appropriate Spike Level

Since the LOD is an estimate for the lower level of the calibration curve, the best spiking level is 1 - 5 times the estimated detection level, as specified in the EPA procedure.

- For ammonia, a concentration between 0.05 and 0.25 mg/L is appropriate.
- For phosphorus, a spike concentration between 0.01 and 0.05 mg/L should result in a valid LOD.

Calculations

LOD = (t-value) (s_x) *With 7 replicates, the t-value is 3.143; for 8 replicates, the t-value to use is 2.998.*

Three important things to remember about calculating LODs are: 1) use the sample standard deviation, 2) use the correct Student's t-value and 3) use all significant figures. The sample standard deviation, s, must be used when calculating LODs. One of the most common mistakes is using the population standard deviation, "s" (*refer to the section on calculating matrix spike control limits for a more detailed explanation*).

Carry all significant figures through the calculations, and round the final LOD to the number of digits used when reporting results for that method. It is acceptable to round the calculated value up to the nearest decimal place. For example, if the calculated LOD is 0.15, it is acceptable to round the LOD to 0.2 if results are only reported to one significant figure. LODs should never be rounded down, unless the laboratory feels it can routinely achieve the rounded value.

Frequency of LOD Determination

Method detection limits will change over time for a variety of reasons, and it is necessary to periodically update the calculated LOD value. Many analytical methods require that the LOD be determined prior to using a new analytical system, and some even require annual updates. The frequency of the determination specified in the analytical method should be followed. If the method does not specify a frequency, the Department recommends that LODs be recalculated whenever a new analyst begins generating data or the performance of the analytical system changes (in addition to the initial LOD). Method detection limits should also be recalculated whenever the analytical procedure is modified (e.g. new extraction solvent). Laboratories may wish to give their LODs an "expiration date" of one year, beyond which the limit is no longer valid, to help maintain current and usable results.

The 5- Point Check

The calculated LOD should be evaluated using several checks to determine if it will meet all of the necessary criteria. The following five actions, which will be collectively referred to as the "Five Point Check", are simple ways to evaluate a calculated LOD.

1. Does the spike level exceed 10 times the LOD? If so, the spike level is high.
2. Is the calculated LOD higher than the spike level? If so, the spike level is too low.
3. Does the calculated LOD meet regulatory requirement (i.e., permit limits)?
4. Is the signal/noise (S/N) in the appropriate range?
5. Are the replicate recoveries reasonable?

Items 1, 2, and 3 are requirements for all LODs. Because even the lowest permit limits are substantially greater than the LODs which can be reasonably achieved for BOD, TSS, ammonia, and phosphorus, item 3 can be ignored for wastewater laboratories whose testing is limited to these parameters. Items 4 and 5 are not required, but are useful for evaluating the LOD data.

Due to the dependence of precision on concentration, the calculated LOD must be greater than one-tenth of the spike level. This is the maximum concentration for an LOD study, and concentrations below this maximum are preferable. At the other extreme, the calculated LOD must not be higher than the spike level. Logically, if the calculated LOD exceeds the spike level it is not statistically possible to differentiate the spiked samples from a blank (and the precision of the determination was very poor!). The following inequalities are useful for evaluating a calculated LOD:

Calculated LOD < **Spike Level** < 10 x Calculated LOD

If these conditions are met the spike level is appropriate. If these conditions are not met, it is necessary to recalculate the LOD. The Department will not accept LOD data if both of the above conditions are not met.

Exempt Analytes

Analytes which are not regulated under Ch. NR 149, including temperature, pH, nutrients in soil and sludge, physical properties of soil and sludges, residual chlorine, specific conductance, flow measurements and microbiological tests are exempt from detection limit calculation requirements.

Residual chlorine is not regulated by the Laboratory Certification Program, but monitoring requirements exist in many permits. There are several approved methods for the determination of residual chlorine including the ion selective electrode (ISE) procedure produced by Orion Research, Inc.. The Iodometric Titration (method 330.2) for residual chlorine analysis is not recommended because it is unlikely to detect residual chlorine in the range required by most wastewater permits. The detection limit for residual chlorine is difficult to calculate due to the unavailability of standards. Department data shows that the LOD can be calculated using the ISE method and household bleach. Matrix specific LODs can also be calculated provided care is used in choosing the appropriate test sample. Laboratories should be capable of achieving an LOD of 0.05 mg/L for residual chlorine.

BOD

BOD detection limits are theoretically based. The assumption is that the **least** amount of depletion allowable is 2 mg/L. Therefore the LOD for a sample in which 300 mL was used is 2 mg/L x (300/300) = 2 mg/L. This technique doesn't involve seed correction, since we are just looking at the overall minimum depletion possible.

Thus, if the highest sample volume used is:	The LOD (for that sample) is:
300 mL	2 mg/L
200 mL	3 mg/L
100 mL	6 mg/L
75 mL	8 mg/L
50 mL	12 mg/L

Example:

Dilution 1	volume used is 100 mL	BOD is	20
Dilution 2	volume used is 75 mL	BOD is	18
Dilution 3	volume used is 50 mL	BOD is	21

Assuming all three dilutions met depletion criteria, you would report a BOD of 19.6 (20 + 18 + 21) ÷ 3) with an LOD of 6 (based on the highest volume used was 100 mL).

If inadequate depletion (i.e., less than 2 mg/L depletion) is obtained on each dilution, using volumes of 150, 100, 50 mL , the lab should (1) Report "< 4" on the DMR, and (2) use MORE sample volume!

TSS

The detection limit for TSS works similarly, but it is based on a minimum capture of 1 mg of residue. Thus, if 1 liter of sample is filtered, then you can "detect" 1 mg per liter. Similarly, if only 500 mL of sample is filtered, then you can "detect" 1mg x (1000/500) = 2 mg/L.

Sample volume filtered	Detection limit
25 mL	40 mg/L
50 mL	20 mg/L
100 mL	10 mg/L
250 mL	4 mg/L

From this, you can see that the real issue for a facility with an effluent permit limit of 10 mg/L or less is they **must** use MORE than 100 mL for their samples in order to get an LOD at or below their permit limit.

NOTE: Even if your plant has a relatively high effluent limit, it is to your advantage to maintain a low LOD since NR 101 fees are calculated based on concentration reported.

Ammonia and Phosphorus

LODs for these parameters should be developed using the standard EPA technique (page 5-1 in the "Yellowbook"). Also consult the document "Analytical Detection Limit Guidance", available from the Central Office.

Ammonia Example

This example calculates the LOD for ammonia-nitrogen using the approved ion selective electrode method. The electrode's manufacturer claims that the probe can detect ammonia as low as 0.05 mg/L. The first step, estimating the LOD, requires familiarity with the analytical procedure. In this case, the manufacturer's claim is a convenient place to start. Multiplying the manufacturer's number, 0.05 mg/L by a factor of 5 gives 0.25 mg/L for a maximum initial spike level (the spike level could be anywhere between 0.05 to 0.25 mg/L). Thus a 0.1 mg/L spike level was selected

Rep. 1	0.104	
Rep. 2	0.082	
Rep. 3	0.096	
Rep. 4	0.1	
Rep. 5	0.087	
Rep. 6	0.114	
Rep. 7	0.108	
mean	**0.098714**	
st dev.	**0.011354**	
t-value	**3.143**	from table based on # replicates
LOD	**0.035684**	t-value x std deviation

# replicates	t-value
7	3.143
8	2.998
9	2.896
10	2.821

The 5-point check

1. Is the LOD no lower than 10% of the spike level? **Yes** [0.035x10] > 0.1

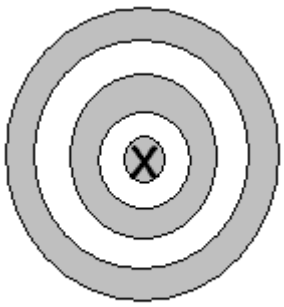
2. Is the spike level greater than the calculated LOD? **Yes** 0.035 < 0.100

3. Is the LOD below any relevant permit limit? **N/A**
(if there is one) Permit limit=

4. Is the signal-to-noise ratio (S/N) between 2.5 and 10? **Yes** 0.098/0.011 = 8.69
S/N = Mean/std dev. S/N= 8.69

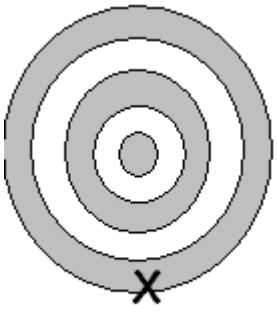
5. Is mean recovery within reasonably expected limits? **Yes** 98+% is excellent
Mean recovery = mean/spike level x 100 98.71%

EVALUATING ACCURACY (MATRIX SPIKES)



Accurate

Accuracy, or bias, as it is termed with increasing frequency, is a measure of systematic error (of an analytical process). Analytical procedures themselves typically have some degree of inherent bias, but the larger component—and that over which the laboratory has some control—is bias resulting from the way the laboratory performs the method. For example, if the laboratory uses automatic pipettors which have been poorly maintained, inaccuracies can develop in the results generated. These inaccuracies are not related to the analytical chemistry, but rather to the laboratories' practices.



Inaccurate

Accuracy is important to both the laboratory and the Department. The Department needs accurate data to minimize adverse effects of discharging an effluent to the biota and water quality of the receiving system. The laboratory has a keen interest in accurate data as well. In the absence of accurate data, wastewater process control becomes extremely difficult to manage, resulting in significantly greater operating costs and labor requirements.

Recovery data obtained from the analysis of spiked samples is used to assess the accuracy of the analysis. Spiking the sample refers to the process of adding a known amount of the target analyte to an aliquot of the sample. Samples are spiked before any preliminary treatment step, such as distillation or digestion. Spiked samples are analyzed at a frequency of 1 per 20 samples analyzed for phosphorus and ammonia. Spikes are not applicable to other analyses performed in the lab.

Tree City laboratory personnel spike the influent at different concentrations than its effluent.

Table 5 illustrates ways to set up spikes for un-distilled ammonia and digested samples for the analysis of phosphorous. Volumetric pipettes are used to measure the spiking solution.

NOTE: The amount of analyte in the spike should be somewhere between 1 and 5 times the known (or expected) concentration of the analyze in the sample which is to be spiked. The analyst may modify this guideline if necessary to stay within the concentration range of the method. If the sample contains no analyte, spike at a level equal to the midrange of the calibration.

TABLE 5 - Preparation of Spikes for Ammonia and Total Phosphorus

PARAMETER + PREPARATION	SAMPLE CONCENTRATION RANGE (MG/L)	STANDARD CONCENTRATION (MG/L)	STANDARD VOLUME (ML)	FINAL VOLUME (ML)	SPIKE CONCENTRATION (MG/L)
Ammonia,	0.0 - 0.5	100	0.5	100	0.50
Un-Distilled	> 0.5 - 1.0	100	1.0	100	1.0
ISE Method	> 1.0 - 2.0	1000	0.2	100	2.0
	> 2.0 - 5.0	1000	0.50	100	5.0
	> 5.0 - 10	1000	1.0	100	10.0
Total	0.00 - 0.25	25.0	0.5	50	0.25
Phosphorus,	> 0.25 - 0.50	25.0	1.0	50	0.50
Digestion	> 0.50 - 2.5	250	0.5	50	2.5
	> 2.5 - 5.0	250	1.0	50	5.0
	> 5.0 - 10.0	250	2.0	50	10.0

If a distillation technique is used before ammonia analysis, the following table can be used:

PARAMETER + PREPARATION	SAMPLE CONCENTRATION RANGE (MG/L)	STANDARD CONCENTRATION (MG/L)	STANDARD VOLUME (ML)	FINAL VOLUME (ML)	SPIKE CONCENTRATION (MG/L)
Ammonia, Distillation	0.0 - 0.5	100	2.0	500	0.40
	> 0.5 - 1.0	100	5.0	500	1.0
	> 1.0 - 5.0	1000	2.5	500	5.0
	> 5.0 - 10	1000	5.0	500	10.0
	> 10	1000	10.0	500	20.0

Accuracy control limits are segregated by matrix. The accuracy of the influent, process control, and effluent are monitored separately. The accuracy limits of the wastewater usually are broader for influent than effluent due to chemical interferences. Influent spike recoveries are summarized separately from effluent spike recoveries. Influent accuracy limits are calculated separately from effluent accuracy limits and influent accuracy control charts are plotted separately from effluent accuracy control charts.

NR 149 requires that quality control limits for spiked and replicates be calculated using a method from an authoritative source. The method described here is considered by many to be the simplest.

✔ **Step 1 - Calculate % Recovery of Individual matrix spikes**
Percent recovery (P) of the spiked amount is calculated by the following equation:

$$P = \frac{\text{observed - background}}{\text{spike amount}} \times 100$$

Where:

observed = the concentration measured in spiked sample (mg/L)
background = the concentration measured in unspiked sample (mg/L)
spike amount = the concentration of spike added to sample (mg/L)

$$\text{spike amount} = \frac{\text{Volume of Standard (ml)} \times \text{Concentration of Standard (mg/L)}}{\text{Final Volume (ml)}}$$

NOTE: For samples which do **not** require digestion of distillation such as ammonia by ion specific electrode, adding the spiking solution increases the volume of spiked sample relative to the unspiked sample. Since Tree City uses the Ion Selective Electrode (ISE) method for analyzing ammonia samples and has an applicable comparison study on file, the facility does not have to distill its samples. However, for those analyses in which digestion is required (Ascorbic Acid method for Total Phosphorus), "background" and "spike" adjustments for volume are not necessary because one spikes before digestion and one always dilutes to the same final volume after digestion.

Unless the dilution resulting from the addition of the spike to the sample is deemed statistically negligible, the background concentration must be adjusted for volume. *If the manner in which spikes are prepared results in a dilution of the sample by 1% or less, the Laboratory Certification and Registration Program considers this dilution to be negligible, and you may directly subtract the background sample concentration without correction.*

background = $\frac{\text{mL of sample}}{\text{mL of sample + mL of spike standard}}$ **X** mg/L measured in unspiked sample

Example:
mL of sample = 48
mL of spike standard = 2
concentration of unspiked sample = 1.50 mg/L
background = 1.50 X $\frac{48}{48 + 2}$ = 1.50 X 0.96 = 1.44 mg/L

The spike concentration is calculated similarly:

spike concentration = $\frac{\text{mL of spike standard}}{\text{mL of spike standard + mL of sample}}$ **X** spike standard concentration (mg/L)

Example:
mL of sample = 48
mL of spike solution = 2
concentration of spike standard = 50 mg/L

spike concentration = 50 X $\frac{2}{48 + 2}$ = 50 X 0.04 = 2.0 mg/L

In this example, if you analyze the **spiked sample** and obtain a **concentration of 3.6 mg/L**:

% Recovery = $\frac{3.6 - 1.44}{2.0}$ X 100 = $\frac{2.16}{2.0}$ X 100 = **108%**

Some operators/analysts have been trained to perform spike-related calculations in terms of the absolute mass (rather than concentration) of the analyte in question. To convert the units used to express the spike amount from mg/L to mg use the following equations:

Spiked sample = mg/L X sample volume (L) = 3.6 mg/L X 0.05 L = 0.180 mg

Background = mg/L X sample volume (L) = 1.44 mg/L X 0.05 L = 0.072 mg

Spike Amount = mL of spike standard X Concentration of Standard (mg/L) X (1 L/1000 mL)
= 2 x 50 X 0.001 = 0.100

Thus spike recovery, expressed as milligrams, = [(1.80 - 0.72) / 0.100] X 100 = 108%

NOTE: Expressing the units in milligrams (mg) may work better for titration methods than for electrode or spectrophotometric methods because the volume of spike added doesn't affect the titration calculation.

✔ **Step 2 - Calculate the average % Recovery (of at least 20 data points)**

The average percent recovery (<P>) is calculated by adding all of the recoveries and dividing by the number of results added. "∑ " is the mathematical symbol indicating summation; thus "∑P" means the sum of all the recoveries. The calculation must have at least 20 results to be statistically valid (and is required by s. NR 149.14 (3)(g), Wis. Adm. Code).

$$< P > = \frac{\sum P}{n}$$

Given the following 20 data points:

	89.3	107.1	95.6	96.3	102.0
	92.7	101.9	103.6	91.6	93.3
	104.3	92.5	97.8	102.2	88.9
	99.5	84.8	100.4	108.4	101.7

Sum (∑P)= 1953.9
data points (n) = 20
Mean Recovery = 97.695

✔ **Step 3 - Calculate the standard deviation of the data set**

This is the formula for sample standard deviation (s_x) of percent recovery.

$$s_x = \sqrt{\frac{\sum\left(P - \langle P \rangle\right)^2}{n-1}}$$

NOTE: another annotation frequently used for standard deviation is "σ".

Using the previous set of data:

P		<P>		P - <P>	$(P - \langle P \rangle)^2$
89.3	-	97.695	=	-8.395	70.476
92.7	-	97.695	=	-4.995	24.950
104.3	-	97.695	=	6.605	43.626
99.5	-	97.695	=	1.805	3.258
107.1	-	97.695	=	9.405	88.454
101.9	-	97.695	=	4.205	17.682
92.5	-	97.695	=	-5.195	26.988
84.8	-	97.695	=	-12.895	166.281
95.6	-	97.695	=	-2.095	4.389
103.6	-	97.695	=	5.905	34.869

P		<P>		P - <P>	$(P - \langle P \rangle)^2$
97.8	-	97.695	=	0.105	0.011
100.4	-	97.695	=	2.705	7.317
96.3	-	97.695	=	-1.395	1.946
91.6	-	97.695	=	-6.095	37.149
102.2	-	97.695	=	4.505	20.295
108.4	-	97.695	=	10.705	114.597
102.0	-	97.695	=	4.305	18.533
93.3	-	97.695	=	-4.395	19.316
88.9	-	97.695	=	-8.795	77.352
101.7	-	97.695	=	4.005	16.040

1. Sum $(P - \langle P \rangle)^2$ =793.5295
2. #1 / (N-1) = 41.7647 (N = 20, so N- 1 = 19)
3. Square root of #2 = 6.4626

Tree City Wastewater Treatment Facility uses a calculator with scientific calculations for these computations. The calculator has two keys for standard deviation, σ_x and s_x. For the above calculation, use the key labeled s_x.

NOTE: Some calculators have σ and σ_{n-1} keys. Use the σ_{n-1} key in this situation.

✔ **Step 4 - Calculate control limits**

Upper Control Limit (UCL) = $\langle P \rangle + 3s_x$
Upper Warning Limit (UWL) = $\langle P \rangle + 2s_x$

Lower Warning Limit (LWL) = $\langle P \rangle - 2s_x$
Lower Control Limit (LCL) = $\langle P \rangle - 3s_x$

NR 149.14(3)(g) states that if, after 12 months, less than 20 spike or replicate QC samples results are generated, control limits may be established based on one of the following:

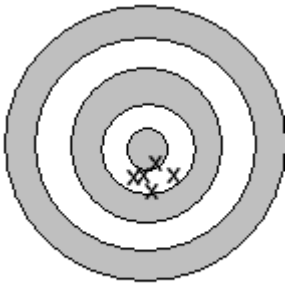
- an authoritative source [such as Standard Methods],

  NOTE: Although this detailed information has been excluded from the19th edition of Standard Methods, Table 1020 I in method 1020B in the 17th edition suggests that matrix spike control limits of 80-120% should be readily achievable for wastewater parameters. Based on the experience of the Laboratory Certification and Registration personnel, these limits are still considered reasonable.

- the experience of the analyst in analyzing samples of this nature, or
- the experience of another laboratory in analyzing a similar sample.

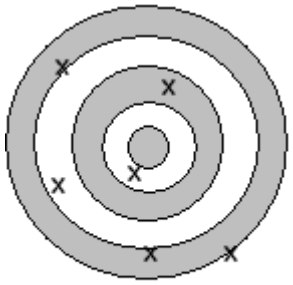
  NOTE: To find out the experience of another lab in analyzing similar samples, one could call a comparable size treatment plant handling the same type of raw influent. If control limits are established in this manner, the specific source and rationale for the selection of limits should be documented.

EVALUATING PRECISION (REPLICATES)



Precise

Precision is a measure of how closely multiple determinations performed on the same sample will agree with each other. To examine precision under the brightest light, consider the implications of analyzing replicates of a final effluent sample in which one of the determinations measures only 80% of a daily maximum permit limit, and the other exceeds the daily maximum by 20%. Which result is correct? One would have an overwhelming desire to report the value that is under the daily maximum, but is that a correct, or even defensible approach? The simple answer is, "No."



Imprecise

Replicate analyses for all requested parameters are performed regularly after every 20 samples. The difference between the results of replicate analyses, the duplicate range, is used to assess the precision (reproducibility) of the test method. Statistical treatment of replicates is as follows:

✔ **Step 1 - Calculate the range (R) of each set of replicate samples**

R = | (Sample result) – (Replicate result) | *where: |expression | = the absolute value of "expression"*

NOTE: The absolute value simply means that negative values are converted to positive values, thus |2 - 5| = 3 rather than -3. Range can be thought of as the difference between two sample results that is always a positive value.

✔ **Step 2 - Calculate the average Range (of at least 20 replicate pairs)**

The average range (<R>) is calculated for ***n*** replicate sample sets (to be statistically significant, *n* must be at least 20):

$$< R > = \frac{\sum R}{n}$$

Sample TSS Effluent Replicate Data

Sample	Replicate	**Range**	Sample	Replicate	**Range**
1.3	1.1	**0.2**	9.3	8.5	**0.8**
2.4	2.5	**0.1**	10.2	9.2	**1**
3.4	3	**0.4**	10.2	10.1	**0.1**
3.6	2.9	**0.7**	11.7	10.7	**1**
3.7	3.3	**0.4**	12.2	11.7	**0.5**
4.1	3.3	**0.8**	14.3	13.2	**1.1**
5.9	5	**0.9**	15.4	15	**0.4**
6	6.1	**0.1**	17.2	16.1	**1.1**
7.5	8.7	**1.2**	17.4	23.2	**5.8**
8	7.8	**0.2**	22.3	21.3	**1**

Sum = 17.8
n = 20
Mean = 0.89

✔ **Step 3 - Calculate control limits**

The control and warning limits are calculated:

Control Limit (CL) = 3.27 <R> for the data set above, CL = 2.91
Warning Limit (WL) = 2.51 <R> for the data set above, WL = 2.23

TABLE 6 - QUALITY CONTROL LIMITS

TEST	KNOWN STANDARDS	REPLICATES			MATRIX SPIKES		
		<R>	Warning Limit	Control Limit	<P>, s	Warning Limits	Control Limits
BOD	198 ± 30.5	0.45	1.1	1.5	NA	NA	NA
TSS	NA	0.6	1.5	2.0	NA	NA	NA
Ammonia-Nitrogen	± 10%	0.05	0.13	0.16	100, 6.67	87 - 113	90 - 110
Total Phosphorus	± 10%	0.07	0.18	0.23	100, 3.33	87 - 107	80 - 120
pH	NA	*0.02*	*0.05*	*0.06*	NA	NA	NA
Chlorine Residual	± 10%				NA	NA	NA
Fecal Coliform	NA				NA	NA	NA

The control limits listed in this table represent limits which can be achieved with the proper training and quality control program. Each lab must set their own limits based on the results of their QC tests. If the limits calculated using actual monitoring data are significantly different than those described in this table, an analytical problem should be suspected. The appropriate corrective action measures should be taken to determine the nature of the problem and make any necessary modifications to procedures to correct the problem. If control limits cannot be improved, refer to DNR district office for further guidance.

USING CONTROL LIMITS APPROPRIATELY

The basic premise behind the practice of updating control limits is to use past data to obtain control limits against which future data are to be evaluated. The single most common mistake made by operators is to take a set of data points, calculate control limits from those data, and then construct a control chart plotting those original data against the control limits established from them.

If you re-calculate control limits annually, then control limits based on the past year's data must be used to evaluate data generated during the next year.

FINE-TUNING CONTROL LIMITS

Occasionally, after calculating new control limits, the limits are perceived to be too restrictive for ordinary work. An example of this might be matrix spike control limits of 96 to 104%. Clearly, there is little room for error in these situations.

The statistical procedure used in the calculation of control limits is designed to produce limits at the 99% confidence level. In layman's terms, this means that only 1 out of every 100 data points is expected to be judged incorrectly (i.e., a point outside of control limits that should really be in-control, or a point within the control limits that should more appropriately be out-of-control). The range of values covered by the mean plus or minus 2 standard deviations is used to obtain warning limits. Statistically speaking, these are 95% confidence limits, which means that 5 out of 100 values (1 out of every 20) is expected to naturally occur outside of these boundaries.

What all of this means is that under normal circumstances, only 1 out of 100 data points is expected to fall outside of control limits—whether that be precision or accuracy. Many laboratories, however, might feel that "out-of-control" data points are found much more frequently. What typically is occurring in these cases however, is that either (1) control limits are not be calculated correctly, (2) there is some inherent bias in the procedure, or (3) some procedural error is at fault.

One of the most common causes behind QC sample failure is that control limits have become unreasonably restrictive. Invariably, this anomaly results from the exclusion of any out-of-control data point from use in constructing future control limits. The misconception is that because the data point is technically "out-of-control", it is not a valid data point, and thus should be excluded. In reality, it is precisely these occasional excursions that provide for a more realistic standard deviation, which in turn affects control limits.

Consider a laboratory with control limits of 80 to 120% and a recovery 0f 121%. Sure, that recovery is an "out-of-control" data point, but why should that point be excluded from calculation of new control limits when a recovery of exactly 120% --the upper limit of control without being "out-of-control"-- would not? It's even easier to see this if you consider a recovery of 120.1% and you assume that control limits of 80 to 120% actually imply 80.0% to 120.0%. One has to remember the distinction between "out-of-control" and "outlier". While outliers should certainly be excluded from calculation of new control limits, the term "outlier" is reserved for those data points which can be statistically defined as such. Simply put then, you MUST include all data points when calculating control limits unless (1) you have documentation which substantiates that an analytical error was made (e.g., you documented that you spilled some of the spike solution or spiked sample during preparation), or (2) you have performed a statistical test which has determined the value to be an outlier.

IDENTIFYING OUTLIER DATA

A number of statistical tests to identify outliers are available, including a procedure used by the U.S.G.S., Dixon's test, the t-test, and Grubbs' test. While any of these tests can be employed to identify outlier data, Grubbs' test will be mentioned here due to its relative simplicity.

First, it is critical to understand that the basic premise of any outlier test is simply to assign a statistical likelihood that a given suspect outlier point represents a different population than the other values in a data set. It cannot –and should not—tell you what to do with that point. Because, however, the consequences associated with failing to exclude outliers (control limits become unreasonably broad allowing inaccurate data) far outweighs the risk of excluding them (control limits become too tight and too many data points become suspect), we recommend that outliers be excluded from further use.

Grubbs' test is based on the calculation of a "Z-score" which is then evaluated against a set of criterion values based on the number of data points in the data set.

$$Z = \frac{|\,(\text{Mean}) - (\text{Suspect result})\,|}{\text{standard deviation}}$$

where: |expression| = absolute value of "expression"

If the resultant Z-score is **greater than** the corresponding criterion value based on the number of data points, then the suspect value is indeed an outlier, and should be excluded and the control limits re-calculated.

Given the following data set:
81, 87, 90, 93, **80**, 82, 91, 94, 83, **125**, 91, 92, 80, 88, 94, 92, 83, 80, 91, 95

You would test both the minimum value (80) and the maximum value (125). The mean for the data set (including both suspect points) is 89.6, and the standard deviation is 9.88. This yields a Z-score of 3.58 for the maximum value (125) and a Z-score of 0.97 for the minimum value (80). The Z-criterion value for a set of 20 data points is 2.71. Consequently, the value, 125, is an outlier. **Control limits resulting from the inclusion of this data point would be 60 to 119%, while if this point is excluded, the control limits would be 71% to 104%.**

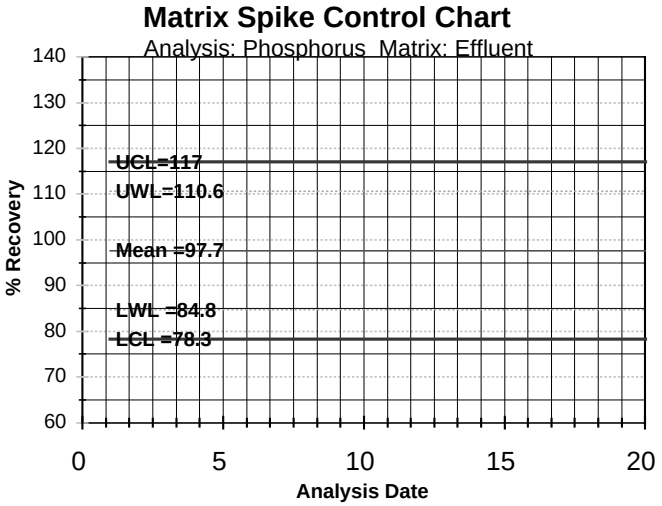
N	Critical Z
15	2.55
16	2.59
17	2.62
18	2.65
19	2.68
20	2.71
21	2.73
22	2.76
23	2.78
24	2.80
25	2.82
26	2.84

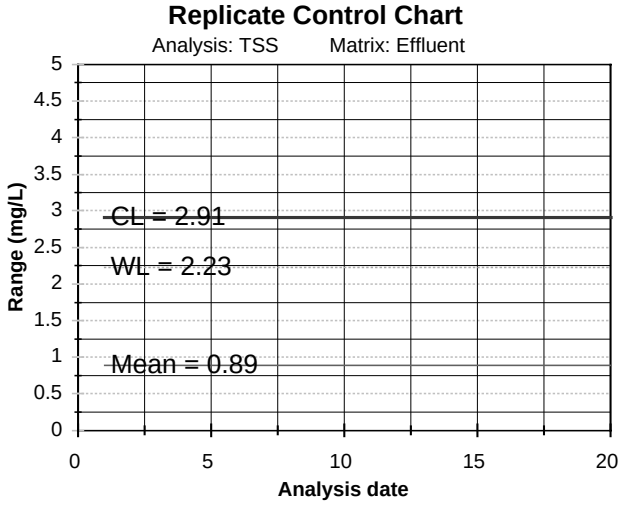
N	Critical Z
27	2.86
28	2.88
29	2.89
30	2.91
31	2.92
32	2.94
33	2.95
34	2.97
35	2.98
36	2.99
37	3.00
38	3.01

N	Critical Z
39	3.03
40	3.04
50	3.13
60	3.20
70	3.26
80	3.31
90	3.35
100	3.38
110	3.42
120	3.44
130	3.47
140	3.49

QUALITY CONTROL CHARTS

Control charts are **not** required by NR 149, but are very useful; it is often much easier to identify the source of analytical problems when a graphic representation of the quality control data (i.e., control charts) is reviewed. Once the chart is constructed, quality control values from future analyses are plotted on it. When 20 new data points have been generated, control limits are again updated. The following sample control charts were generated using the data from the previous examples.





It is important to remember that the quality control results used to generate a control chart are NOT plotted on that chart. Those limits should be used to construct a new, blank, QC chart-- which is filled in over time as new data are generated.

The laboratory must also keep a detailed record of historical quality control limits, as this information will be reviewed during an NR 149 on-site evaluation. One of the components of an audit involves tracing data back during the prior three-year compliance period. The auditor will be looking for the laboratory's ability to produce the quality control limits associated with past data. With this information, the auditor can determine whether any data exceeded control limits, and, if so, what corrective action was taken in response to the exceedance.

Precision control limits and charts must, at times, be further segregated by concentration. Influent replicates are summarized separately from effluent replicates because influent samples typically contain significantly greater concentrations of analytes than the final, effluent. Influent precision limits are calculated separately from effluent limits. (See Table 6). Influent control charts are plotted separately from effluent control charts.

EVALUATING QUALITY CONTROL CHARTS

Quality Control charts provide a valuable tool for the analyst to "see the big picture". For example, it's often difficult to initially assess the cause of a matrix spike that exceeded control limits. By looking at a control chart, however, the analyst may be better able to identify a trend of decreasing recoveries that eventually led to the exceedance. The goal, then, is to use control charts as a form of preventive medicine. Routine review of control charts will enable the analyst to detect –and correct—a potential analytical problem before it leads to an exceedance.

The following conditions indicate an out-of-control situation. Analysts should be monitoring their control charts for the development of any of these trends. When any QC exceedance is encountered, the operator/analyst must take corrective action, as directed by s. NR 149.14 (3)(h), Wis. Adm. Code.

- Any point beyond the upper or lower control limits
- Seven successive points on the same side of the <P> value.
- Another criterion frequently used to identify out-of-control situations is a trend of 5 or more data points moving in the same direction (i.e. 5 points, each of which is successfully higher (or lower) than the previous one).

The 19th edition of Standard Methods suggests the following additional situations warrant corrective action:

- Repeat analyses exceeding either control limit. If any measurement exceeds a control limit, repeat the QC analysis once. If the repeat analysis falls within limits proceed, otherwise stop and correct the problem before proceeding.
- 3 of 4 successive data points outside warning limits
- 4 of 5 successive data points exceeding the mean $\pm$ 1 standard deviation
- 8 successive analyses above (or below) the mean

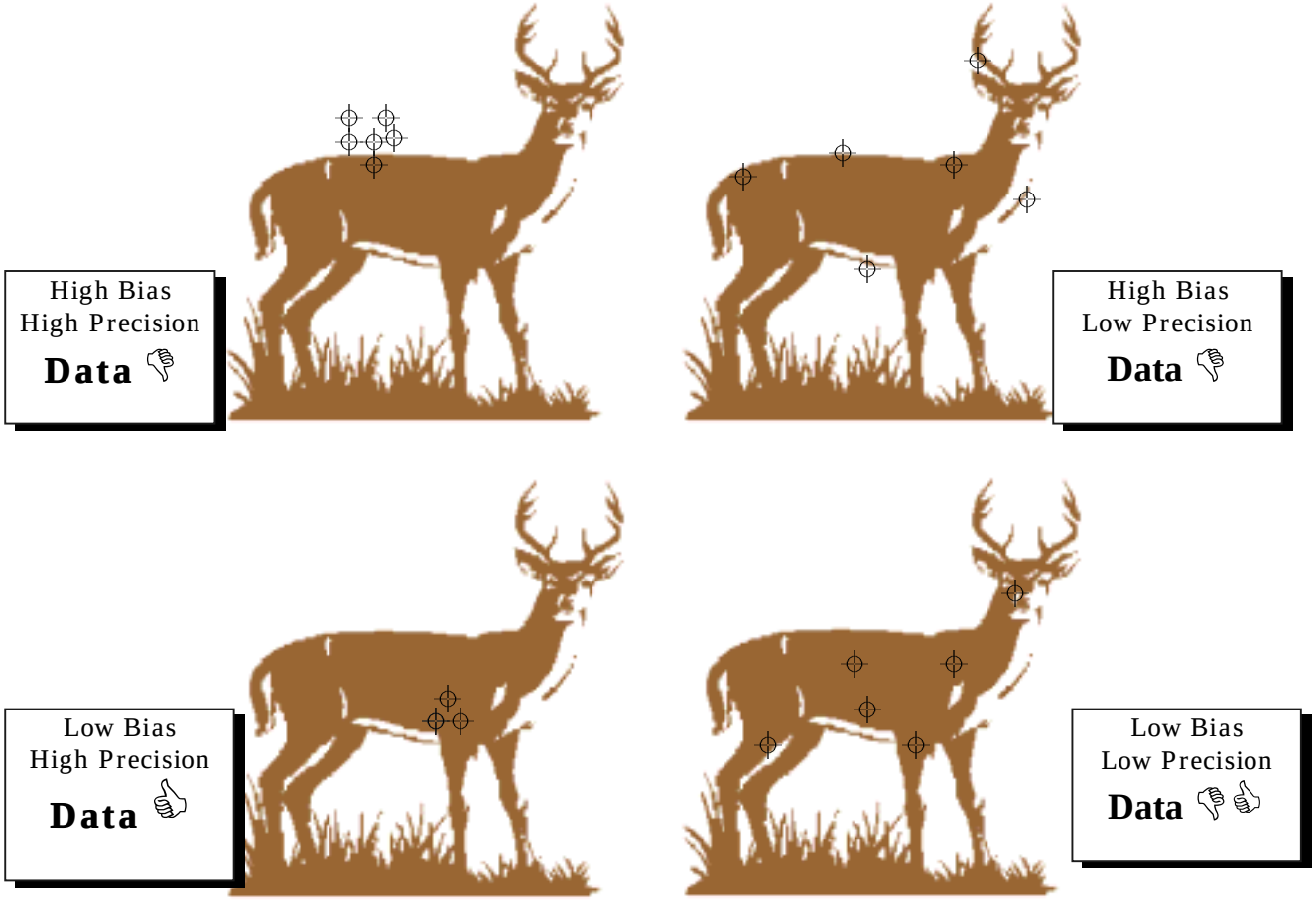
If any of these conditions occur, lab priorities should be directed to identifying and resolving the problem. One of the best ways to quickly resolve an out-of-control event, while minimizing the amount of data which must be qualified, is to analyze replicates or matrix spikes for each sample run until the out-of-control situation ceases. The problem and solutions are documented; all results since the last in-control point are reanalyzed or are qualified.

Providing a sufficient number of quality control results are generated, new control limits should be calculated every six months, using data produced during the previous six month period and new control charts are generated. The Lab Certification Code does not specify how often control limits should be updated. The six month frequency given here is a suggestion only. While appropriate for a facility with Tree City's analysis schedule, a longer interval may be necessary for very small facilities.

Control limits for precision must be segregated by both matrix and concentration, when necessary. Influent replicates are summarized in separate records from effluent replicates. Influent precision limits are calculated independently from effluent precision limits (See Table 6). Influent precision control charts are constructed separately from effluent precision control charts. Precision data must be further segregated by concentration if a concentration dependency exists. For example, if the influent concentration varies greatly, then some WWTPs may have to further segregate their influent TSS into "low" and "high" concentrations. A good general division is that concentration equal to or greater than ten (10) times the limit of detection (LOD). You may need to contact your regional certification officer for further guidance if this is your case.

RELATING BIAS & PRECISION TO DATA QUALITY

As the following figures illustrate, both accuracy (low bias) and precision are required to generate valid data.



CORRECTIVE ACTION

To put it simply, corrective action is required whenever any control limit is exceeded. The goal of corrective action is to develop a historical record which details the type of problems a laboratory encounters as well as the resolution process itself. To put this into perspective, most operators who have analyzed BOD using a DO probe come to realize that the DO probe membrane degrades over time, and at some certain point, you need to replace it. Most of this information, unfortunately, is kept in the operator's head. A new operator—perhaps an individual without extensive analytical experience—who takes over laboratory responsibilities, may not recognize the "warning signs" that indicate a need to change the membrane. Without a corrective action trail, which essentially can be reduced to documentation in the form of "*When you see this, do the following*", analysis will continue, until the fateful day that a GGA sample fails to meet acceptance criteria. Once data has been generated under circumstances in which one or more QC checks exceed their associated acceptance criteria, it is too late, and further effort is mandatory.

Corrective action, then, should be considered to be a sort of preventive medicine against future illnesses. We should learn from our past "difficulties" to guard against going through the same difficulties in the future. As an example, the first time one ever "messes" with a bee's nest, they tend to get stung. The best corrective action for this situation might simply be to stay away from bee's nests. Is that the only corrective action for this situation? Certainly not! The kid that chooses a corrective action of simply throwing the rock (at the bee's nest) from further away may find, however, that his "corrective action" was not adequate to correct the "problem".

A laboratory needs to document the results obtained from all corrective action efforts so that in the future, the operator/analyst can "home in" on the right solution for a particular problem. A good example of this is to consider what corrective action might be taken in response to failing a matrix spike control limit for recovery. In this example, let's say an operator records a matrix spike recovery of 52% for total phosphorus. Being in a hurry, as many things need to be done around the village, the operator's corrective action is to review his calculations to make sure he didn't make a simple math error. Satisfied that no error was made with calculations, he shrugs his shoulders, chalks it up to fate, and goes about his other duties. This same process goes on for several weeks, and over that time, he gets recoveries between 42% and 58%. He's heard about matrix interferences, and figures that his effluent has suddenly developed a matrix interference. Six months later, however, it's time for an evaluation of his laboratory and the DNR auditor discovers that the real problem is that the standard used to prepare matrix spikes not only should have been prepared fresh 3 months ago, but that an error was made during the preparation. The spike solution is found to be exactly one-half of the concentration the operator believed it to be.

Remember...we need to be sure that the corrective action taken in response to a QC failure is appropriate for the problem at hand, and the laboratory needs to establish a system to check back and ensure that the problem has actually been solved. After all, if one brings their car into the garage to have it looked at because it won't start, they would not be happy to hear that the mechanic's "corrective action" included performing a costly re-alignment!

Last resort - Qualifying your results

Operators will still need to report those analytical results that are associated with any analytical run in which one or more of the quality control samples fails to meet acceptance criteria. These data must be flagged on the DMR reports by putting a check on the Quality Control line in the upper right hand corner of the form as well as by circling or marking the affected data with an asterisk. It is also acceptable to include a narrative that describes which dates and analyses are affected, and specific details regarding the reason for qualification of the data. The operator must also decide whether or not to include the analytical results when calculating weekly or monthly average values. If the decision is made to exclude the values in question from calculating weekly/monthly averages, an explanation for the exclusion(s) must also be provided.

DOCUMENTATION AND RECORD-KEEPING

All records of equipment calibration and maintenance, QC tests, sampling, and sample analysis are retained for at least three (3) years [five (5) years for sludge data] at the treatment facility office in metal file cabinets. Before any result is reported, all raw data and calculations are reviewed for accuracy by someone other than the analyst/lab technician who performed the testing. Depending on the size of the lab/wastewater facility, the individual who performs the review on analytical testing may be another analyst/lab technician, operator or the facility supervisor. In any case, the individual must have sufficient experience to be capable of distinguishing between correct and incorrect data.

It is important that all raw data be kept, no matter how rough in appearance. This information can be very helpful in locating data handling problems. If data contained on any record is transcribed to facilitate summarizing or neatness, the original record must also be kept. Sections NR 149.06 (1) (a) and (b) specifically require that records be maintained such that *"...any sample may be traced back to the analyst, date collected, date analyzed, and method used including raw data, intermediate calculations, results, and the final report*." and that "*Quality control results shall be traceable to all of the associated sample results.*"

A typical wastewater treatment plant operator is charged with a multitude of duties throughout the municipality, and when something has to be overlooked, it tends to be documentation. Unfortunately, documentation is essential to substantiating the quality of your data. A good principle to operate by is that if you did the work – whatever that aspect of the analysis might be—take credit for your efforts, by documenting what you did. The old adage, "*If you didn't document it...you didn't do it*", also applies.

Finally, s. NR 149.06 (5) requires that all *records "shall be handled in a manner to ensure their permanence and security*." This section of NR 149 also specifies that "*Handwritten records shall be recorded in ink*." In keeping with the intent of these requirements, correction fluid of any type, and pens containing "eraseable ink" must not be used. If errors are made when making a notation, a single line should be drawn through the entry and the correction entered directly above. Good laboratory practice also requires that corrections made in this manner should be accompanied by the initials of the individual making the correction and the date the correction was made.